File Copy

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, A61K 38/17

A2

(11) International Publication Number: **WO 98/56909**

(43) International Publication Date: 17 December 1998 (17.12.98)

(21) International Application Number: PCT/US98/11822

(22) International Filing Date: 8 June 1998 (08.06.98)

(30) Priority Data:

08/873,218	11 June 1997 (11.06.97)	US
09/092,722	5 June 1998 (05.06.98)	US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 113 Ann Lee Road, Harvard, MA 01451 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US). HOWES, Steven, H.; Apartment 2, 44 Chester Street, Somerville, MA 02144 (US). FECHTEL, Kim; 46 Marion Road, Arlington, MA 02174 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Polynucleotides and the proteins encoded thereby are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

This application is a continuation-in-part of application Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/873,218), filed June 11, 1997, which is incorporated by reference herein.

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 12 to nucleotide 800;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 78 to nucleotide 800;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1 to nucleotide 547;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bh389_11 deposited under accession number ATCC 98451;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bh389_11 deposited under accession number ATCC 98451;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bh389_11 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bh389_11 deposited under accession number ATCC 98451;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:2;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 12 to nucleotide 800; the nucleotide sequence of SEQ ID NO:1 from

nucleotide 78 to nucleotide 800; the nucleotide sequence of SEQ ID NO:1 from nucleotide 1 to nucleotide 547; the nucleotide sequence of the full-length protein coding sequence of clone bh389_11 deposited under accession number ATCC 98451; or the nucleotide sequence of a mature protein coding sequence of clone bh389_11 deposited under accession number ATCC 98451. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bh389_11 deposited under accession number ATCC 98451. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 178. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:2, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 126 to amino acid 135 of SEQ ID NO:2.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 178;

(c) fragments of the amino acid sequence of SEQ ID NO:2 comprising eight consecutive amino acids of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone bh389_11 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 178. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID

NO:2, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2having biological activity, the fragment comprising the amino acid sequence from amino acid 126 to amino acid 135 of SEQ ID NO:2.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 100 to nucleotide 882;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 635 to nucleotide 867;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bk112_15 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bk112_15 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bk112_15 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bk112_15 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 100 to nucleotide 882; the nucleotide sequence of SEQ ID NO:3

from nucleotide 635 to nucleotide 867; the nucleotide sequence of the full-length protein coding sequence of clone bk112_15 deposited under accession number ATCC 98451; or the nucleotide sequence of a mature protein coding sequence of clone bk112_15 deposited under accession number ATCC 98451. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bk112_15 deposited under accession number ATCC 98451. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 200 to amino acid 256. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:4, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 125 to amino acid 134 of SEQ ID NO:4.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 200 to amino acid 256;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising eight consecutive amino acids of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone bk112_15 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 200 to amino acid 256. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:4, or a protein comprising a fragment of the amino acid sequence of

SEQ ID NO:4having biological activity, the fragment comprising the amino acid sequence from amino acid 125 to amino acid 134 of SEQ ID NO:4.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 245 to nucleotide 520;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 181 to nucleotide 527;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bk200_13 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bk200_13 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bk200_13 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bk200_13 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:6;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 245 to nucleotide 520; the nucleotide sequence of SEQ ID NO:5 from nucleotide 181 to nucleotide 527; the nucleotide sequence of the full-length protein

coding sequence of clone bk200_13 deposited under accession number ATCC 98451; or the nucleotide sequence of a mature protein coding sequence of clone bk200_13 deposited under accession number ATCC 98451. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bk200_13 deposited under accession number ATCC 98451. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:6, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 41 to amino acid 50 of SEQ ID NO:6.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) fragments of the amino acid sequence of SEQ ID NO:6 comprising eight consecutive amino acids of SEQ ID NO:6; and

(c) the amino acid sequence encoded by the cDNA insert of clone bk200_13 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:6, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 41 to amino acid 50 of SEQ ID NO:6.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 365 to nucleotide 784;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 518 to nucleotide 784;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone di386_3 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone di386_3 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone di386_3 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone di386_3 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:8;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 365 to nucleotide 784; the nucleotide sequence of SEQ ID NO:7 from nucleotide 518 to nucleotide 784; the nucleotide sequence of the full-length protein coding sequence of clone di386_3 deposited under accession number ATCC 98451; or the nucleotide sequence of a mature protein coding sequence of clone di386_3 deposited under accession number ATCC 98451. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone di386_3 deposited under accession number ATCC 98451. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein

comprising the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 140. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:8, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 65 to amino acid 74 of SEQ ID NO:8.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:7 or SEQ ID NO:9.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 140;

(c) fragments of the amino acid sequence of SEQ ID NO:8 comprising eight consecutive amino acids of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone di386_3 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8 or the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 140. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:8, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8having biological activity, the fragment comprising the amino acid sequence from amino acid 65 to amino acid 74 of SEQ ID NO:8.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 191 to nucleotide 781;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 56 to nucleotide 492;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone em397_2 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone em397_2 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone em397_2 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone em397_2 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:11;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:10 from nucleotide 191 to nucleotide 781; the nucleotide sequence of SEQ ID NO:10 from nucleotide 56 to nucleotide 492; the nucleotide sequence of the full-length protein coding sequence of clone em397_2 deposited under accession number ATCC 98451; or the nucleotide sequence of a mature protein coding sequence of clone em397_2 deposited under accession number ATCC 98451. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone em397_2 deposited under accession number ATCC 98451. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein

comprising the amino acid sequence of SEQ ID NO:11 from amino acid 1 to amino acid 101. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:11, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity, the fragment comprising the amino acid sequence from amino acid 93 to amino acid 102 of SEQ ID NO:11.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:10.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:11;

(b) the amino acid sequence of SEQ ID NO:11 from amino acid 1 to amino acid 101;

(c) fragments of the amino acid sequence of SEQ ID NO:11 comprising eight consecutive amino acids of SEQ ID NO:11; and

(d) the amino acid sequence encoded by the cDNA insert of clone em397_2 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:11 or the amino acid sequence of SEQ ID NO:11 from amino acid 1 to amino acid 101. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:11, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11having biological activity, the fragment comprising the amino acid sequence from amino acid 93 to amino acid 102 of SEQ ID NO:11.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 65 to nucleotide 1636;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 482 to nucleotide 1636;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 487 to nucleotide 1006;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fh170_7 deposited under accession number ATCC 98451;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fh170_7 deposited under accession number ATCC 98451;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fh170_7 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fh170_7 deposited under accession number ATCC 98451;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:13;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:13 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:13;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:12 from nucleotide 65 to nucleotide 1636; the nucleotide sequence of SEQ ID NO:12 from nucleotide 482 to nucleotide 1636; the nucleotide sequence of SEQ ID NO:12 from nucleotide 487 to nucleotide 1006; the nucleotide sequence of the full-length protein coding sequence of clone fh170_7 deposited under accession number ATCC 98451; or the nucleotide sequence of a mature protein coding sequence of clone fh170_7 deposited under accession number ATCC 98451. In other preferred embodiments, the

polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fh170_7 deposited under accession number ATCC 98451. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:13 from amino acid 142 to amino acid 314. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:13 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:13, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:13 having biological activity, the fragment comprising the amino acid sequence from amino acid 257 to amino acid 266 of SEQ ID NO:13.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:12.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:13;

(b) the amino acid sequence of SEQ ID NO:13 from amino acid 142 to amino acid 314;

(c) fragments of the amino acid sequence of SEQ ID NO:13 comprising eight consecutive amino acids of SEQ ID NO:13; and

(d) the amino acid sequence encoded by the cDNA insert of clone fh170_7 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:13 or the amino acid sequence of SEQ ID NO:13 from amino acid 142 to amino acid 314. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:13 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:13, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:13having biological activity, the fragment comprising the amino acid sequence from amino acid 257 to amino acid 266 of SEQ ID NO:13.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 41 to nucleotide 550;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fn53_4 deposited under accession number ATCC 98451;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fn53_4 deposited under accession number ATCC 98451;

(e) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fn53_4 deposited under accession number ATCC 98451;

(f) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fn53_4 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:16;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(h).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:15 from nucleotide 41 to nucleotide 550; the nucleotide sequence of the full-length protein coding sequence of clone fn53_4 deposited under accession number ATCC 98451; or the nucleotide sequence of a mature protein coding sequence of clone fn53_4 deposited under accession number ATCC 98451. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fn53_4 deposited under accession number ATCC 98451. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16 from amino acid 40 to amino acid

170. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:16, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising the amino acid sequence from amino acid 80 to amino acid 89 of SEQ ID NO:16.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:15, SEQ ID NO:14 or SEQ ID NO:17 .

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 40 to amino acid 170;

(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising eight consecutive amino acids of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone fn53_4 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:16 or the amino acid sequence of SEQ ID NO:16 from amino acid 40 to amino acid 170. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:16, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising the amino acid sequence from amino acid 80 to amino acid 89 of SEQ ID NO:16.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 84 to nucleotide 404;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 78 to nucleotide 493;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fq505_4 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fq505_4 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fq505_4 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fq505_4 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:19;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:19;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:18 from nucleotide 84 to nucleotide 404; the nucleotide sequence of SEQ ID NO:18 from nucleotide 78 to nucleotide 493; the nucleotide sequence of the full-length protein coding sequence of clone fq505_4 deposited under accession number ATCC 98451; or the nucleotide sequence of a mature protein coding sequence of clone fq505_4 deposited under accession number ATCC 98451. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fq505_4 deposited under accession number ATCC 98451. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:19 from amino acid 23 to amino acid 107. In further preferred embodiments, the present invention provides a polynucleotide

encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:19, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 having biological activity, the fragment comprising the amino acid sequence from amino acid 48 to amino acid 57 of SEQ ID NO:19.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:18.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:19;

(b) the amino acid sequence of SEQ ID NO:19 from amino acid 23 to amino acid 107;

(c) fragments of the amino acid sequence of SEQ ID NO:19 comprising eight consecutive amino acids of SEQ ID NO:19; and

(d) the amino acid sequence encoded by the cDNA insert of clone fq505_4 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:19 or the amino acid sequence of SEQ ID NO:19 from amino acid 23 to amino acid 107. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:19, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19having biological activity, the fragment comprising the amino acid sequence from amino acid 48 to amino acid 57 of SEQ ID NO:19.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 1439 to nucleotide 1744;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 1241 to nucleotide 1754;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fw13_9 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fw13_9 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fw13_9 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fw13_9 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:21;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:20 from nucleotide 1439 to nucleotide 1744; the nucleotide sequence of SEQ ID NO:20 from nucleotide 1241 to nucleotide 1754; the nucleotide sequence of the full-length protein coding sequence of clone fw13_9 deposited under accession number ATCC 98451; or the nucleotide sequence of a mature protein coding sequence of clone fw13_9 deposited under accession number ATCC 98451. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fw13_9 deposited under accession number ATCC 98451. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 57. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein

comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:21, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity, the fragment comprising the amino acid sequence from amino acid 46 to amino acid 55 of SEQ ID NO:21.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:20.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:21;

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 57;

(c) fragments of the amino acid sequence of SEQ ID NO:21 comprising eight consecutive amino acids of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone fw13_9 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:21 or the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 57. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:21, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21having biological activity, the fragment comprising the amino acid sequence from amino acid 46 to amino acid 55 of SEQ ID NO:21.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 47 to nucleotide 919;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 124 to nucleotide 452;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone gg619_2 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone gg619_2 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone gg619_2 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone gg619_2 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:23;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:22 from nucleotide 47 to nucleotide 919; the nucleotide sequence of SEQ ID NO:22 from nucleotide 124 to nucleotide 452; the nucleotide sequence of the full-length protein coding sequence of clone gg619_2 deposited under accession number ATCC 98451; or the nucleotide sequence of a mature protein coding sequence of clone gg619_2 deposited under accession number ATCC 98451. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone gg619_2 deposited under accession number ATCC 98451. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23 from amino acid 27 to amino acid 135. In further preferred embodiments, the present invention provides a polynucleotide

encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:23, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23 having biological activity, the fragment comprising the amino acid sequence from amino acid 140 to amino acid 149 of SEQ ID NO:23.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:22.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:23;

(b) the amino acid sequence of SEQ ID NO:23 from amino acid 27 to amino acid 135;

(c) fragments of the amino acid sequence of SEQ ID NO:23 comprising eight consecutive amino acids of SEQ ID NO:23; and

(d) the amino acid sequence encoded by the cDNA insert of clone gg619_2 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:23 or the amino acid sequence of SEQ ID NO:23 from amino acid 27 to amino acid 135. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:23, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23having biological activity, the fragment comprising the amino acid sequence from amino acid 140 to amino acid 149 of SEQ ID NO:23.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

(a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and

(b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature forms) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g. , receptors) from the cell

in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Clone "bh389_11"

A polynucleotide of the present invention has been identified as clone "bh389_11". bh389_11 was isolated from a human adult thyroid cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bh389_11 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bh389_11 protein").

The nucleotide sequence of bh389_11 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bh389_11 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Amino acids 10 to 22 of SEQ ID NO:2 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 23, or are a transmembrane domain. The TopPredII computer program predicts a potential transmembrane domain within the bh389_11 protein sequence centered around amino acid 68 of SEQ ID NO:2.

Another potential bh389_11 reading frame and predicted amino acid sequence is encoded by basepairs 757 to 1833 of SEQ ID NO:1 and is reported in SEQ ID NO:34. A frameshift in the nucleotide sequence of SEQ ID NO:1 between about nucleotide 754 to about nucleotide 803 could join the reading frames of SEQ ID NO:1 and SEQ ID NO:34. The TopPredII computer program predicts a potential transmembrane domain within the amino acid sequence of SEQ ID NO:34, centered around amino acid 357 of SEQ ID NO:34.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bh389_11 should be approximately 1700 bp.

The nucleotide sequence disclosed herein for bh389_11 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bh389_11 demonstrated at least some similarity with sequences identified as AA307880 (EST178733 Colon carcinoma (HCC) cell line Homo sapiens cDNA 5' end), AA442426 (zv70f06.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 759011 5'), H70103 (yr92f04.r1 Homo sapiens cDNA clone 212767 5'), R19820

(yg37f12.r1 Homo sapiens cDNA clone 34771 5'), and W46238 (zc30e10.s1 Soares senescent fibroblasts NbHSF Homo sapiens cDNA clone 323850 3'). Based upon sequence similarity, bh389_11 proteins and each similar protein or peptide may share at least some activity.

Clone "bk112_15"

A polynucleotide of the present invention has been identified as clone "bk112_15". bk112_15 was isolated from a human adult retina cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bk112_15 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bk112_15 protein").

The nucleotide sequence of bk112_15 as presently determined is reported in SEQ ID NO:3. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bk112_15 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bk112_15 should be approximately 1300 bp.

The nucleotide sequence disclosed herein for bk112_15 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bk112_15 demonstrated at least some similarity with sequences identified as AA307119 (EST178031 Colon carcinoma (HCC) cell line Homo sapiens cDNA 5' end), AA318352 (EST20422 Retina II Homo sapiens cDNA 5' end similar to similar to C. elegans hypothetical protein, cosmid ZK688.2), L20941 (Human ferritin heavy chain mRNA, complete cds), M97164 (Human ferritin heavy chain mRNA, complete cds), N25339 (yx55d08.s1 Homo sapiens cDNA clone 265647 3'), N31453 (yx55d08.r1 Homo sapiens cDNA clone 265647 5'), and N33227 (yy07d02.s1 Homo sapiens cDNA clone 270531 3' similar to gb:L20941 FERRITIN HEAVY CHAIN (HUMAN)). The predicted amino acid sequence disclosed herein for bk112_15 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted bk112_15 protein demonstrated at least some similarity to sequences identified as Z68335 (C29F4.2 [Caenorhabditis elegans]). Based

upon sequence similarity, bk112_15 proteins and each similar protein or peptide may share at least some activity.

Clone "bk200_13"

A polynucleotide of the present invention has been identified as clone "bk200_13". bk200_13 was isolated from a human adult retina cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bk200_13 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bk200_13 protein").

The nucleotide sequence of bk200_13 as presently determined is reported in SEQ ID NO:5. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bk200_13 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bk200_13 should be approximately 1000 bp.

The nucleotide sequence disclosed herein for bk200_13 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bk200_13 demonstrated at least some similarity with sequences identified as AA098915 zk84f06.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 489539 3'), AA150367 zl07b06.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 491603 5'), AA235904 (zs40h05.r1 Soares NhHMPu S1 Homo sapiens cDNA clone 687705 5'), N32487 (yx79g10.r1 Homo sapiens cDNA clone 268002 5'), and T47862 (yb17g03.r1 Homo sapiens cDNA clone 71476 5'). Based upon sequence similarity, bk200_13 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of bk200_13 may contain CAAAAA repeat-like elements.

Clone "di386_3"

A polynucleotide of the present invention has been identified as clone "di386_3". di386_3 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was

identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. di386_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "di386_3 protein").

The nucleotide sequence of the 5' portion of di386_3 as presently determined is reported in SEQ ID NO:7. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:8. The predicted amino acid sequence of the di386_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8. Amino acids 39 to 51 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 52, or are a transmembrane domain. Amino acids 17 to 29 OF SEQ ID NO:8 are also a possible leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 30, or are a transmembrane domain. Additional nucleotide sequence from the 3' portion of di386_3, including the polyA tail, is reported in SEQ ID NO:9.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone di386_3 should be approximately 2000 bp.

The nucleotide sequence disclosed herein for di386_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. di386_3 demonstrated no similarity with any known sequences in those databases.

Clone "em397_2"

A polynucleotide of the present invention has been identified as clone "em397_2". em397_2 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. em397_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "em397_2 protein").

The nucleotide sequence of em397_2 as presently determined is reported in SEQ ID NO:10. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the em397_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:11.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone em397_2 should be approximately 1250 bp.

The nucleotide sequence disclosed herein for em397_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. em397_2 demonstrated at least some similarity with sequences identified as AA092876 (m0851.seq.F Fetal heart, Lambda ZAP Express Homo sapiens cDNA 5'), AA180952 (zp41b06.r1 Stratagene muscle 937209 Homo sapiens cDNA clone 611987 5'), AA463323 (zx71f01.r1 Soares total fetus Nb2HF8 9w Homo sapiens), H87081 (ys74f01.r1 Homo sapiens cDNA clone 220537 5'), W56381 (zc57a01.r1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone 326376 5'), W88527 (zh73g02.s1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA clone 417746 3'), and Z64565 (H.sapiens CpG island DNA genomic Mse1 fragment, clone 13d12, reverse read cpg13d12.rt1c). Based upon sequence similarity, em397_2 proteins and each similar protein or peptide may share at least some activity.

Clone "fh170_7"

A polynucleotide of the present invention has been identified as clone "fh170_7". fh170_7 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. fh170_7 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fh170_7 protein").

The nucleotide sequence of fh170_7 as presently determined is reported in SEQ ID NO:12. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the fh170_7 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:13. Amino acids 127 to 139 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 140, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fh170_7 should be approximately 2200 bp.

The nucleotide sequence disclosed herein for fh170_7 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and

FASTA search protocols. fh170_7 demonstrated at least some similarity with sequences identified as AA112479 (zn69a02.s1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone 563402 3'), AA593402 (nn57g10.s1 NCI_CGAP_Kid6 Homo sapiens cDNA clone IMAGE:1088034), Q76795 (Human genome fragment), T26136 (Human gene signature HUMGS08373), and Z19759 (H. sapiens putatively transcribed partial sequence; UK-HGMP sequence ID AAAALWX; single read). The predicted amino acid sequence disclosed herein for fh170_7 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted fh170_7 protein demonstrated at least some similarity to sequences identified as D32253 (MagA [Magnetospirillum sp.]) and W01520 (MagA protein). The predicted fh170_7 protein also demonstrated at least some similarity to other prokaryotic membrane transport proteins: potassium-efflux system protein kefB and NaH-antiporter protein. Based upon sequence similarity, fh170_7 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts ten potential transmembrane domains within the fh170_7 protein sequence, centered around amino acids 130, 160, 210, 230, 280, 310, 360, 380, 420, and 500 of SEQ ID NO:13, respectively.

Clone "fn53_4"

A polynucleotide of the present invention has been identified as clone "fn53_4". fn53_4 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. fn53_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fn53_4 protein").

The nucleotide sequence of the 5' portion of fn53_4 as presently determined is reported in SEQ ID NO:14. An additional internal nucleotide sequence from fn53_4 as presently determined is reported in SEQ ID NO:15. What applicants believe is the proper reading frame and the predicted amino acid sequence encoded by such internal sequence is reported in SEQ ID NO:16. Additional nucleotide sequence from the 3' portion of fn53_4, including the polyA tail, is reported in SEQ ID NO:17.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fn53_4 should be approximately 4100 bp.

The nucleotide sequence disclosed herein for fn53_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. fn53_4 demonstrated at least some similarity with sequences identified as AA179207 (zp46c11.s1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone 612500 3'), AA279207 (zs83e06.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:704098 3', mRNA sequence), H87151 (yw15a06.s1 Homo sapiens cDNA clone 252274 3'), and H83373 (ys90a09.r1 Homo sapiens cDNA clone 222040 5' similar to SP:BICD_DROME P16568 CYTOSKELETON-LIKE BICAUDAL D). The predicted amino acid sequence disclosed herein for fn53_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted fn53_4 protein demonstrated at least some similarity to sequences identified as M31684 and X51652 (bicaudalD protein [Drosophila melanogaster]) and R66930 (AMML chromosome inv(16) product). Based upon sequence similarity, fn53_4 proteins and each similar protein or peptide may share at least some activity.

Clone "fq505_4"

A polynucleotide of the present invention has been identified as clone "fq505_4". fq505_4 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. fq505_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fq505_4 protein").

The nucleotide sequence of fq505_4 as presently determined is reported in SEQ ID NO:18. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the fq505_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:19.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fq505_4 should be approximately 512 bp.

The nucleotide sequence disclosed herein for fq505_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. fq505_4 demonstrated at least some similarity with sequences identified as Z71861 (C.hircus mRNA for EST2-31). The predicted amino acid sequence

disclosed herein for fq505_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted fq505_4 protein demonstrated at least some similarity to sequences identified as P92141 (Recombinant human adult T cell leukaemia derived factor polypeptide), X54539 (thioredoxin [Homo sapiens]), and X77584 (ATL-derived factor/thioredoxin [Homo sapiens]). The predicted fq505_4 protein also demonstrated at least some similarity to sequences identified as surface associated sulphydryl protein (GenProt accession number135773). The similarity between these proteins includes a WCGPC catalytic site, which is present as RCGPC at amino acids 31 to 35 of the predicted fq505_4 protein. In addition to having thioredoxin catalytic activity, at least one thioredoxin-related protein has also been reported to be "an IL-2 receptor/Tac inducer" (Tagaya *et al.*, 1989, *EMBO J.* **8(3)**: 757-764). At least one thioredoxin-related protein is reported to be associated with the plasma membrane, "indicating that the protein may be a member of this [thioredoxin] family and function as an essential growth factor" (Martin and Dean, 1991, *Biochem. Biophys. Res. Commun.* **175(1)**: 123-128). Based upon sequence similarity, fq505_4 proteins and each similar protein or peptide may share at least some activity.

Clone "fw13_9"

A polynucleotide of the present invention has been identified as clone "fw13_9". fw13_9 was isolated from a human adult testes (teratocarcinoma NCCIT) cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. fw13_9 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fw13_9 protein").

The nucleotide sequence of fw13_9 as presently determined is reported in SEQ ID NO:20. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the fw13_9 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:21.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fw13_9 should be approximately 1900 bp.

The nucleotide sequence disclosed herein for fw13_9 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and

FASTA search protocols. fw13_9 demonstrated at least some similarity with sequences identified as AA047557 (zf13f08.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 376839 5'), AA284524 (zt20d07.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 713677 3'), AA502778 (ne43e04.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:900126), J04743 (M.musculus Ms6-hm locus, repeat elements), R35040 (yh86a10.r1 Homo sapiens cDNA clone 136602 5'), T21414 (Human gene signature HUMGS02783), and U91318 (Human chromosome 16p13 BAC clone CIT987SK-962B4 complete sequence). Based upon sequence similarity, fw13_9 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the fw13_9 protein sequence centered around amino acid 30 of SEQ ID NO:21.

Clone "gg619_2"

A polynucleotide of the present invention has been identified as clone "gg619_2". gg619_2 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. gg619_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "gg619_2 protein").

The nucleotide sequence of gg619_2 as presently determined is reported in SEQ ID NO:22. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the gg619_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:23.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone gg619_2 should be approximately 1350 bp.

The nucleotide sequence disclosed herein for gg619_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. gg619_2 demonstrated at least some similarity with sequences identified as N42957 (yy12b12.r1 Homo sapiens cDNA clone 271007 5' similar to SW:ALG5_YEAST P40350 dolichyl-phosphate beta-glucosyltransferase), N50844 (yy91g05.s1 Homo sapiens cDNA clone 280952 3' similar to SW:ALG5_YEAST P40350 dolichyl-phosphate beta-glucosyltransferase), and N62597 (yz75a06.s1 Homo sapiens

cDNA clone 288850 3' similar to SW:ALG5_YEAST P40350 Dolichyl-phosphate beta-glucosyltransferase). The predicted amino acid sequence disclosed herein for gg619_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted gg619_2 protein demonstrated at least some similarity to sequences identified as R38093 (nodC N-terminal portion [Bradyrhizobium sp. (Parasponia)]) and X77573 (dolichyl-phosphate beta-glucosyl-transferase [Saccharomyces cerevisiae]). The enzyme UDP-glucose:dolichyl-phosphate glucosyltransferase is a transmembrane-bound enzyme of the endoplasmic reticulum involved in protein N-linked glycosylation, and catalyzes the transfer of glucose from UDP-glucose to dolichyl phosphate. Based upon sequence similarity, gg619_2 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the gg619_2 protein sequence, centered around amino acid 188 of SEQ ID NO:23.

Deposit of Clones

Clones bh389_11, bk112_15, bk200_13, di386_3, em397_2, fh170_7, fn53_4, fq505_4, fw13_9, and gg619_2 were deposited on June 10, 1997 with the American Type Culture Collection (10801 University Boulevard, Manassas, Virginia 20110-2209 U.S.A.) as an original deposit under the Budapest Treaty and were given the accession number ATCC 98451, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b), and the term of the deposit will comply with 37 C.F.R. § 1.806.

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Figures 1A and 1B, respectively. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* **19**: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* **9**: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped"

(i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of an oligonucleotide probe that was used to isolate or to sequence each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

Clone	Probe Sequence
bh389_11	SEQ ID NO:24
bk112_15	SEQ ID NO:25
bk200_13	SEQ ID NO:26
di386_3	SEQ ID NO:27
em397_2	SEQ ID NO:28
fh170_7	SEQ ID NO:29
fn53_4	SEQ ID NO:30
fq505_4	SEQ ID NO:31
fw13_9	SEQ ID NO:32
gg619_2	SEQ ID NO:33

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoaramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramadite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

(a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;

(b) It should be designed to have a T_m of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with g-^{32}P ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately 4e+6 dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 µl of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 µg/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 µg/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to 1e+6 dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S. McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form(s) of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence(s) of the mature form(s) of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that

has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* **15(7)**: 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* **62(1)**: 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* **58**: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* **14(9)**: 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* **90(16)**: 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* **91(2)**: 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* **336**: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614, 396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that

the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologues of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide. Preferably, polynucleotide species homologues have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, and protein species homologues have at least 30% sequence identity (more preferably, at least 45% identity; most preferably at least 60% identity) with the given protein, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides or the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Species homologues may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. Preferably, species homologues are those isolated from mammalian species. Most preferably, species homologues are those isolated from certain mammalian species such as, for example, *Pan troglodytes, Gorilla gorilla, Pongo pygmaeus, Hylobates concolor, Macaca mulatta, Papio papio, Papio hamadryas, Cercopithecus aethiops, Cebus capucinus, Aotus trivirgatus, Sanguinus oedipus, Microcebus murinus, Mus musculus, Rattus norvegicus, Cricetulus griseus, Felis catus, Mustela vison, Canis familiaris, Oryctolagus cuniculus, Bos taurus,*

Ovis aries, *Sus scrofa*, and *Equus caballus*, for which genetic maps have been created allowing the identification of syntenic relationships between the genomic organization of genes in one species and the genomic organization of the related genes in another species (O'Brien and Seuánez, 1988, *Ann. Rev. Genet.* 22: 323-351; O'Brien *et al.*, 1993, *Nature Genetics* 3:103-112; Johansson *et al.*, 1995, *Genomics* 25: 682-690; Lyons *et al.*, 1997, *Nature Genetics* 15: 47-56; O'Brien *et al.*, 1997, *Trends in Genetics* 13(10): 393-399; Carver and Stubbs, 1997, *Genome Research* 7:1123-1137; all of which are incorporated by reference herein).

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotides which also encode proteins which are identical or have significantly similar sequences to those encoded by the disclosed polynucleotides. Preferably, allelic variants have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps. Allelic variants may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from individuals of the appropriate species.

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides that hybridize under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer†
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T_B*; 1xSSC	T_B*; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T_D*; 1xSSC	T_D*; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T_F*; 1xSSC	T_F*; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T_H*; 4xSSC	T_H*; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T_J*; 4xSSC	T_J*; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T_L*; 2xSSC	T_L*; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T_N*; 6xSSC	T_N*; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T_P*; 6xSSC	T_P*; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T_R*; 4xSSC	T_R*; 4xSSC

‡: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

†: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*T_B - T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6($\log_{10}$[Na^+]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na^+] is the concentration of sodium ions in the hybridization buffer ([Na^+] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant

methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, those described in Gyuris *et al.*, 1993, *Cell* **75**: 791-803 and in Rossi *et al.*, 1997, *Proc. Natl. Acad. Sci. USA* **94**: 8405-8410, all of which are incorporated by reference herein) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine

levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is

evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ, Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science *257*:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, *89*:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.

Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter

7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent

myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells.* R.I. Freshney, *et al.* eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells.* R.I. Freshney, *et al.* eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells.* R.I. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc.., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells.* R.I. Freshney, *et al.* eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland,

H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of

congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunolgobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be

administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred

pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1ng to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal

antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-

aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in

the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.
Howes, Steven H.
Fechtel, Kim

(ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(iii) NUMBER OF SEQUENCES: 34

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Genetics Institute, Inc.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: U.S.A.
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Sprunger, Suzanne A.
(B) REGISTRATION NUMBER: 41,323

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 498-8284
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2043 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTGAATGCCC CATGCGCACC CCACAGCTCG CGCTCCTGCA AGTGTTCTTT CTGGTGTTCC 60

CCGATGGCGT CCGGCCTCAG CCCTCTTCCT CCCCATCAGG GGCAGTGCCC ACGTCTTTGG 120

AGCTGCAGCG AGGGACGGAT GGCGGAACCC TCCAGTCCCC TTCAGAGGCG ACTGCAACTC 180

GCCCGGCCGT GCCTGGACTC CCTACAGTGG TCCCTACTCT CGTGACTCCC TCGGCCCCTG 240

GGAATAGGAC TGTGGACCTC TTCCCAGTCT TACCGATCTG TGTCTGTGAC TTGACTCCTG 300

GAGCCTGCGA TATAAATTGC TGCTGCGACA GGGACTGCTA TCTTCTCCAT CCGAGGACAG 360

TTTTCTCCTT CTGCCTTCCA GGCAGCGTAA GGTCTTCAAG CTGGGTTTGT GTAGACAACT 420

CTGTTATCTT CAGGAGTAAT TCCCCGTTTC CTTCAAGAGT TTTCATGGAT TCTAATGGAA 480

TCAGGCAGTT TTGTGTCCAT GTGAACAACT CAAACTTAAA CTATTTCCAG AAGCTTCAAA 540

AGGTCAATGC AACCAACTTC CAGGCCCTGG TTGCAGAGTT TGGAGGCGAA TCATTCACTT 600

CAACATTCCA AACTCAATCA CCACCATCTT TTTACAGGGC CGGGGACCCC ATTCTTACTT 660

ACTTCCCCAA GTGGTCTGTA ATAAGCTTGC TGAGACAACC TGCAGGAGTT GGAGCTGGGG 720

GACTCTGTGC TGAAAGCAAT CCTGCAGGTT TCCTAGAGAG TAAAAGTACA ACTTGCACTC 780

GTTTTTTTCA AGAACCTGGC TAGTAGCTGT ACCTTGGATT CAGCCCTCAA TGCTGCCTCT 840

TACTATAACT TCACAGTCTT AAAGGTTCCA AGAAGCATGA CTGATCCACA GAATATGGAG 900

TTCCAGGTTC CTGTAATACT TACCTCACAG GCTAATGCTC CTCTGTTGGC TGGAAACACT 960

TGTCAGAATG TAGTTTCTCA GGTCACCTAT GAGATAGAGA CCAATGGGAC TTTTGGAATC 1020

CAGAAAGTTT CTGTCAGTTT GGGACAAACC AACCTGACTG TTGAGCCAGG CGCTTCCTTA 1080

CAGCAACACT TCATCCTTCG CTTCAGGGCT TTTCAACAGA GCACAGCTGC TTCTCTCACC 1140

AGTCCTAGAA GTGGGAATCC TGGCTATATA GTTGGGAAGC CACTCTTGGC TCTGACTGAT 1200

GATATAAGTT ACTCAATGAC CCTCTTACAG AGCCAGGGTA ATGGAAGTTG CTCTGTTAAA 1260

AGACATGAAG TGCAGTTTGG AGTGAATGCA ATATCTGGAT GCAAGCTCAG GTTGAAGAAG 1320

GCAGACTGCA GCCACTTGCA GCAGGAGATT TATCAGACTC TTCATGGAAG GCCCAGACCA 1380

GAGTATGTTG CCATCTTTGG TAATGCTGAC CCAGCCCAGA AAGGAGGGTG GACCAGGATC 1440

CTCAACAGGC ACTGCAGCAT TTCAGCTATA AACTGTACTT CCTGCTGTCT CATACCAGTT 1500

TCCCTGGAGA TCCAGGTATT GTGGGCATAT GTAGGTCTCC TGTCCAACCC GCAAGCTCAT 1560

GTATCAGGAG TTCGATTCCT ATACCAGTGC CAGTCTATAC AGGATTCTCA GCAAGTTACA 1620

GAAGTATCTT TGACAACTCT TGTGAACTTT GTGGACATTA CCCAGAAGCC ACAGCCTCCA 1680

AGGGGCCAAC CCAAAATGGA CTGGAAATGG CCATTCGACT TCTTTCCCTT CAAAGTGGCA 1740

TTCAGCAGAG GAGTATTCTC TCAAAAATGC TCAGTCTCTC CCATCCTTAT CCTGTGCCTC 1800

TTAGAACTTG GAGTTCTCAA CCTAGAGACT ATGTGAAGAA AAGAAAATAA TCAGATTTCA 1860

GTTTTCCCTA TGAGAAACTC TGAGGCAGCC ACTTATCTTG GCTAAATAGA ACCTCACCTG 1920

CTCATGACCA GAGAGCATTT AGGATAATAG AGGACCTAAC TGAAGGAATC CTTGTATATG 1980

AAAGGAGTTA TTTTAGAAAA GCAATAAAAA TATTTTATTC ATCATAAAAA AAAAAAAAAA 2040

AAA 2043

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 263 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Arg Thr Pro Gln Leu Ala Leu Leu Gln Val Phe Phe Leu Val Phe
1 5 10 15

Pro Asp Gly Val Arg Pro Gln Pro Ser Ser Ser Pro Ser Gly Ala Val
20 25 30

Pro Thr Ser Leu Glu Leu Gln Arg Gly Thr Asp Gly Gly Thr Leu Gln
35 40 45

Ser Pro Ser Glu Ala Thr Ala Thr Arg Pro Ala Val Pro Gly Leu Pro
50 55 60

Thr Val Val Pro Thr Leu Val Thr Pro Ser Ala Pro Gly Asn Arg Thr
65 70 75 80

Val Asp Leu Phe Pro Val Leu Pro Ile Cys Val Cys Asp Leu Thr Pro
85 90 95

Gly Ala Cys Asp Ile Asn Cys Cys Cys Asp Arg Asp Cys Tyr Leu Leu
100 105 110

His Pro Arg Thr Val Phe Ser Phe Cys Leu Pro Gly Ser Val Arg Ser
115 120 125

Ser Ser Trp Val Cys Val Asp Asn Ser Val Ile Phe Arg Ser Asn Ser
130 135 140

Pro Phe Pro Ser Arg Val Phe Met Asp Ser Asn Gly Ile Arg Gln Phe
145 150 155 160

Cys Val His Val Asn Asn Ser Asn Leu Asn Tyr Phe Gln Lys Leu Gln
165 170 175

Lys Val Asn Ala Thr Asn Phe Gln Ala Leu Val Ala Glu Phe Gly Gly
180 185 190

Glu Ser Phe Thr Ser Thr Phe Gln Thr Gln Ser Pro Pro Ser Phe Tyr
195 200 205

Arg Ala Gly Asp Pro Ile Leu Thr Tyr Phe Pro Lys Trp Ser Val Ile
210 215 220

Ser Leu Leu Arg Gln Pro Ala Gly Val Gly Ala Gly Gly Leu Cys Ala
225 230 235 240

Glu Ser Asn Pro Ala Gly Phe Leu Glu Ser Lys Ser Thr Thr Cys Thr
245 250 255

Arg Phe Phe Gln Glu Pro Gly
260

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1263 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAGCAGCTCA TCAACCCCTT TGGAGAGGAT GATGATGATT TTGAGACCAA CTGGATTGTC 60

GACAGGAATT TGCAGGTGTC CCTGTTGGCT GTGGATGAGA TGCACCAGGA CCTGCCTCGG 120

ATGGAGCCGG ACATGTACTG GAATAAGCCC GAGCCACAGC CCCCCTACAC AGCTGCTTCC 180

GCCCAGTTCC GTCGAGCCTC CTTTATGGGC TCCACCTTCA ACATCAGCCT GAACAAAGAG 240

GAGATGGAGT TCCAGCCCAA TCAGGAGGAC GAGGAGGATG CTCACGCTGG CATCATTGGC 300

CGCTTCCTAG GCCTGCAGTC CCATGATCAC CATCCTCCCA GGGCAAACTC AAGGACCAAA 360

CTACTGTGGC CCAAGAGGGA ATCCCTTCTC CACGAGGGCC TGCCCAAAAA CCACAAGGCA 420

GCCAAACAGA ACGTTAGGGG CCAGGAAGAC AACAAGGCCT GGAAGCTTAA GGCTGTGGAC 480

GCCTTCAAGT CTGCCCCACT GTATCAGAGG CCAGGCTACT ACAGTGCCCC ACAGACGCCC 540

CTCAGCCCCA CTCCCATGTT CTTCCCCCTA GAACCATCAG CGCCGTCAAA GCTTCACAGT 600

GTCACAGGCA TAGACACCAA AGACAAAAGC TTAAAGACTG TGAGTTCTGG GGCCAAGAAA 660

AGTTTTGAAT TGCTCTCAGA GAGCGATGGG GCCTTGATGG AGCACCCAGA AGTATCTCAA 720

GTGAGGAGGA AAACTGTGGA GTTTAACCTG ACGGATATGC CAGAGATCCC CGAAAATCAC 780

CTCAAAGAAC CTTTGGAACA ATCACCAACC AACATACACA CTACACTCAA AGATCACATG 840

GATCCTTATT GGGCCTTGGA AAACAGGGAT GAAGCACATT CCTAACCTGC TTCCTAATGG 900

GGATGCTTCG CCAGCCAGGT CCTCACCTGT GTGTACACCA GCAGGACACT GATCCAGTCA 960

CAGCCATACA GCTGTCCACA CTGAAGAACA TGTCCTACAA CAGCCTGAAT CAAATGGCTA 1020

GCTTAATAGA TAAAAATCCC AGACTACTTC AGCCTTTAAT GCCTTTTATT CATAAAAACT 1080

GTGAAAGCTA GACTGAACCA TTGGAAACAT TTAACTCAGA CTCTGGATTC AGAGTCGGGA 1140

ACCCTTAGTT CTATCTGAAT CCAAGACAGC CACACCTTAG TATACTGCCC AAACTAATGA 1200

GTTTAATAAA TACAAATACT CGTTAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1260

AAA 1263

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 261 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met His Gln Asp Leu Pro Arg Met Glu Pro Asp Met Tyr Trp Asn Lys
1 5 10 15

Pro Glu Pro Gln Pro Pro Tyr Thr Ala Ala Ser Ala Gln Phe Arg Arg

20 25 30

Ala Ser Phe Met Gly Ser Thr Phe Asn Ile Ser Leu Asn Lys Glu Glu
35 40 45

Met Glu Phe Gln Pro Asn Gln Glu Asp Glu Glu Asp Ala His Ala Gly
50 55 60

Ile Ile Gly Arg Phe Leu Gly Leu Gln Ser His Asp His His Pro Pro
65 70 75 80

Arg Ala Asn Ser Arg Thr Lys Leu Leu Trp Pro Lys Arg Glu Ser Leu
85 90 95

Leu His Glu Gly Leu Pro Lys Asn His Lys Ala Ala Lys Gln Asn Val
100 105 110

Arg Gly Gln Glu Asp Asn Lys Ala Trp Lys Leu Lys Ala Val Asp Ala
115 120 125

Phe Lys Ser Ala Pro Leu Tyr Gln Arg Pro Gly Tyr Tyr Ser Ala Pro
130 135 140

Gln Thr Pro Leu Ser Pro Thr Pro Met Phe Phe Pro Leu Glu Pro Ser
145 150 155 160

Ala Pro Ser Lys Leu His Ser Val Thr Gly Ile Asp Thr Lys Asp Lys
165 170 175

Ser Leu Lys Thr Val Ser Ser Gly Ala Lys Lys Ser Phe Glu Leu Leu
180 185 190

Ser Glu Ser Asp Gly Ala Leu Met Glu His Pro Glu Val Ser Gln Val
195 200 205

Arg Arg Lys Thr Val Glu Phe Asn Leu Thr Asp Met Pro Glu Ile Pro
210 215 220

Glu Asn His Leu Lys Glu Pro Leu Glu Gln Ser Pro Thr Asn Ile His
225 230 235 240

Thr Thr Leu Lys Asp His Met Asp Pro Tyr Trp Ala Leu Glu Asn Arg
245 250 255

Asp Glu Ala His Ser
260

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 894 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTTTGAGGGT TTTTTGTTTT TTGTTTTTTC TAGGATTTCA TTGTGATGTT TTGGTTTTGT 60

TTTTTGCTTT TTGTTTAAGT TGTGCTGACA CCAAACACAT CCAGTTTATA ATCAGTACAT 120

TGGAAAGCTG GTATTGATGT AGAACCAGTG CATAACTTTT TATGGGGTTT TGTTATTGGT 180

TTTTTTTTTG TAAAGTGTGA ATAAAAGGTA TGTTTACTCA TTTTTCCTGA ACACTGTGTT 240

GGTAATGTGC ATCATGACAA TTTCCAGTGA AGGTGAGCTG GAGCTGGTTG GACTAATGAG 300

ACTGAGGAAG CAGCTTTTCC TACGATCTGC ATTATGTAAT CACAGGTCCA GAGAGCTTTA 360

TGGAAGCGGG AGAGGAGGAG CACTTACTCA TGTTGTATTT GTTAATGGAG GATGTCATCT 420

TTTCATAGAT GCTGGAACTA GAGTGCACTT GTTAGATGCT AAAGGTTTGA GCTTTACACA 480

AAATGTCTTC ATCTGTATTT GTTATTGTCT ACAATATATT TGAATTTGGG GCAGCATATT 540

AAGATGTAAT GCCCTGTTAT GTCTGGAAAA AACTTGTTTT GCTTCTTCCA GGCAAAGGGC 600

ATTTTGTGGA TCAGTTTGAA CAGCTTCTCC ACCTTATTTG GACAGTGATA AATTGAACCA 660

AGAGTGTAGA TTTACAAGTG TAACCTTCAA AAGAGGAAGA ACTATTTGGG GTCTGTAGGT 720

AATGAACAGT CACACCAAAA TAGACTATGA TGCTTTTGTT AAGAAAGGTT TCATGTTTTA 780

GATATTTTCC GTGTCCTAAA TAATTTTCAA TAATCTATAA TCCCTAAAAT GCAATAAAAA 840

CTAGTATGTT TTCAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAA 894

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 92 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Cys Ile Met Thr Ile Ser Ser Glu Gly Glu Leu Glu Leu Val Gly
1 5 10 15

Leu Met Arg Leu Arg Lys Gln Leu Phe Leu Arg Ser Ala Leu Cys Asn
20 25 30

His Arg Ser Arg Glu Leu Tyr Gly Ser Gly Arg Gly Gly Ala Leu Thr
35 40 45

His Val Val Phe Val Asn Gly Gly Cys His Leu Phe Ile Asp Ala Gly
50 55 60

Thr Arg Val His Leu Leu Asp Ala Lys Gly Leu Ser Phe Thr Gln Asn
65 70 75 80

Val Phe Ile Cys Ile Cys Tyr Cys Leu Gln Tyr Ile
85 90

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 784 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCGGCCGCAG GTCTACTTGT GGCGAGCAGT CCAGCACAGC CTCACAGTGC AGAGCATGAG 60

CTTTGGAGCC TGCCCCCACC CTAGCTTTGT GACCTTAAGT GAGCTACATA GCTTCTCATG 120

TGTAAACTAC TCATCATAAT GGTTCTGACC TCAGTGGTTT GTTGTGTTCT AGGAAATGAT 180

GCCAGTGAAT GCGTAGTCCC AGCCTCAGCA CAGGGGAGCC ACCTTGAAGC TCTCAAATAT 240

CACTGTTGTG AATACAGAGA GGGAAAACCA ACTGTAACGT GCCACCCAAA TTTTTTCAGA 300

TTAATACATC ATTCATCAGA CTTCATTCGT GATCTCGAAG AGTGACATCA GTCTTCCTTG 360

GAATATGAAG AGAATTTCTT TGGTTCTTCT TTTGCATTTC TATTTGATTT ATTTTATTTT 420

ATTTTATTTT ATGTTTTTTG GTACAGAAAG CTCATTACTA GTCCTGTCCA GCAACGTGCC 480

TCTCCTGGCC CTAGAGTTCT TGGAAATAGC CCAGGCCAAA GAGAAGGCCT TTCTCCCCAT 540

GGTCAGCCAC ACGTTCCACA TGCGCACAGA GGAGTCTGAT GCCTCACAGG AGGGCGATGA 600

CCTACCCAAG TCCTCAGCAA ACACCAGCCA TCCCAAGCAG GATGACAGCC CCAAGTCCTC 660

AGAAGAAACC ATCCAGCCCA AGGAGGGTGA CATCCCCAAG GCCCCAGAAG AAACCATCCA 720

ATCCAAGAAG GAGGACCTCC CCAAGTCCTC GGAAAAAGCC ATCCAGCCCA AAGAGAGTAA 780

CATC 784

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 140 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Lys Arg Ile Ser Leu Val Leu Leu Leu His Phe Tyr Leu Ile Tyr
1 5 10 15

Phe Ile Leu Phe Tyr Phe Met Phe Phe Gly Thr Glu Ser Ser Leu Leu
20 25 30

Val Leu Ser Ser Asn Val Pro Leu Leu Ala Leu Glu Phe Leu Glu Ile
35 40 45

Ala Gln Ala Lys Glu Lys Ala Phe Leu Pro Met Val Ser His Thr Phe
50 55 60

His Met Arg Thr Glu Glu Ser Asp Ala Ser Gln Glu Gly Asp Asp Leu
65 70 75 80

Pro Lys Ser Ser Ala Asn Thr Ser His Pro Lys Gln Asp Asp Ser Pro
85 90 95

Lys Ser Ser Glu Glu Thr Ile Gln Pro Lys Glu Gly Asp Ile Pro Lys
100 105 110

Ala Pro Glu Glu Thr Ile Gln Ser Lys Lys Glu Asp Leu Pro Lys Ser
115 120 125

Ser Glu Lys Ala Ile Gln Pro Lys Glu Ser Asn Ile
130 135 140

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 75 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 60

AAAAAAAAAA AAAAA 75

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 939 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGAAGAAGTA GAAGCATCGA AAGCGTTGGA GAGGTGTTAC CGGAACGGCG GCGACAAGGG 60

TGTTCCCGAA CTAGAGTGGG GCATACATAA TCTTGCTGCT ATGCTTCGAA GCTGTAGTCT 120

GAATCAACCT AAGTTTTAAA CAGAAGGTGA ACCTCTGAGA TAGAAAATCA AGTATATTTT 180

AAAAGAAGGG ATGTGGGATC AAGGAGGACA GCCTTGGCAG CAGTGGCCCT TGAACCAGCA 240

ACAATGGATG CAGTCATTCC AGCACCAACA GGATCCAAGC CAGATTGATT GGGCTGCATT 300

GGCCCAAGCT TGGATTGCCC AAAGAGAAGC TTCAGGACAG CAAAGCATGG TAGAACAACC 360

ACCAGGAATG ATGCCAAATG GACAAGATAT GTCTACAATG GAATCTGGTC CAAACAATCA 420

TGGGAATTTC CAAGGGGATT CAAACTTCAA CAGAATGTGG CAACCAGAAT GGGGAATGCA 480

TCAGCAACCC CCACACCCCC CTCCAGATCA GCCATGGATG CCACCAACAC CAGGCCCAAT 540

GGACATTGTT CCTCCTTCTG AAGACAGCAA CAGTCAGGAC AGTGGGGAAT TTGCCCCTGA 600

CAACAGGCAT ATATTTAACC AGAACAATCA CAACTTTGGT GGACCACCCG ATAATTTTGC 660

AGTGGGGCCA GTGAACCAGT TTGACTATCA GGACCTCCAG GACCTCCAGC ACCTCCCCAG 720

AATCGAAGAG AAAGGCCATC ATCATTCAGG GATCGTCAGC GTTCACCTAT TGCACTTCCT 780

GTGAAGCAGG AGCCTCCACA AATTGACGCA GTAAAACGCA GGACTCTTCC CGCTTGGATT 840

CGCGAAGGTC TTGAAAAAAT GGAACGTGAA AAGCAGAAGA AATTGGAGAA AGAAAGAATG 900

GAACAACAAC GTTCACAATT GTCCAAAAAA AAAAAAAAA 939

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 197 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
Met Trp Asp Gln Gly Gly Gln Pro Trp Gln Gln Trp Pro Leu Asn Gln
1               5                  10                  15

Gln Gln Trp Met Gln Ser Phe Gln His Gln Gln Asp Pro Ser Gln Ile
            20                  25                  30

Asp Trp Ala Ala Leu Ala Gln Ala Trp Ile Ala Gln Arg Glu Ala Ser
        35                  40                  45

Gly Gln Gln Ser Met Val Glu Gln Pro Pro Gly Met Met Pro Asn Gly
    50                  55                  60

Gln Asp Met Ser Thr Met Glu Ser Gly Pro Asn Asn His Gly Asn Phe
65                  70                  75                  80

Gln Gly Asp Ser Asn Phe Asn Arg Met Trp Gln Pro Glu Trp Gly Met
                85                  90                  95

His Gln Gln Pro Pro His Pro Pro Pro Asp Gln Pro Trp Met Pro Pro
            100                 105                 110

Thr Pro Gly Pro Met Asp Ile Val Pro Pro Ser Glu Asp Ser Asn Ser
        115                 120                 125

Gln Asp Ser Gly Glu Phe Ala Pro Asp Asn Arg His Ile Phe Asn Gln
    130                 135                 140

Asn Asn His Asn Phe Gly Gly Pro Pro Asp Asn Phe Ala Val Gly Pro
145                 150                 155                 160

Val Asn Gln Phe Asp Tyr Gln Asp Leu Gln Asp Leu Gln His Leu Pro
                165                 170                 175

Arg Ile Glu Glu Lys Gly His His His Ser Gly Ile Val Ser Val His
            180                 185                 190

Leu Leu His Phe Leu
        195
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2343 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AGGAGAGCAG CCGGCAGCGC CTGGAGGCCC TGAGAGAGCT GCAATAAAGG AAGAAACAGA 60

ATATATGGAA CTTCTGGCAG CAGAAAAACA TCAAGTTGAA GCCCTTAAAA ATATGCAACA 120

TCAAAACCAA AGTTTATCCA TGCTTGACGA GATTCTTGAA GATGTAAGAA AGGCAGCGGA 180

TCGTCTGGAG GAAGAGATAG AGGAACATGC TTTTGACGAC AATAAATCAG TCAAGGGGGT 240

CAATTTTGAG GCAGTTCTGA GGGTGGAGGA AGAAGAGGCC AATTCTAAGC AAAATATAAC 300

AAAACGAGAA GTGGAGGATG ACTTGGGTCT TAGCATGCTG ATTGACTCCC AGAACAACCA 360

GTATATTTTG ACCAAGCCCA GAGATTCAAC CATCCCACGT GCAGATCACC ACTTTATAAA 420

GGACATTGTT ACCATAGGAA TGCTGTCCTT GCCTTGTGGC TGGCTATGTA CAGCCATAGG 480

ATTGCCTACA ATGTTTGGTT ATATTATTTG TGGTGTACTT CTGGGACCTT CAGGACTAAA 540

TAGTATTAAG TCTATTGTGC AAGTGGAGAC ATTAGGAGAA TTTGGGGTGT TTTTTACTCT 600

TTTTCTTGTT GGCTTAGAAT TTTCTCCAGA AAAGCTAAGA AAGGTGTGGA AGATTTCCTT 660

ACAAGGGCCG TGTTACATGA CACTGTTAAT GATTGCATTT GGCTTGCTGT GGGGGCATCT 720

CTTGCGGATC AAACCCACGC AGAGCGTCTT CATTTCCACG TGTCTGTCCT TGTCAAGCAC 780

ACCCCTCGTG TCCAGGTTCC TCATGGGCAG TGCTCGGGGT GACAAAGAAG GCGACATTGA 840

CTACAGCACC GTGCTCCTCG GCATGCTGGT GACGCAGGAC GTGCAGCTCG GGCTCTTCAT 900

GGCCGTCATG CCGACTCTCA TACAGGCGGG CGCCAGTGCA TCTTCTAGCA TTGTCGTGGA 960

AGTTCTCCGA ATCCTGGTTT TGATTGGTCA GATTCTTTTT TCACTAGCGG CGGTTTTTCT 1020

TTTATGTCTT GTTATAAAGA AGTATCTCAT TGGACCCTAT TATCGGAAGC TGCACATGGA 1080

AAGCAAGGGG AACAAAGAAA TCCTGATCTT GGGAATATCT GCCTTTATCT TCTTAATGTT 1140

AACGGTCACG GAGCTGCTGG ACGTCTCCAT GGAGCTGGGC TGTTTCCTGG CTGGAGCGCT 1200

CGTCTCCTCT CAGGGCCCCG TGGTCACCGA GGAGATCGCC ACCTCCATCG AACCCATCCG 1260

CGACTTCCTG GCCATCGTTT TCTTCGCCTC CATAGGGCTC CACGTGTTCC CCACGTTTGT 1320

GGCGTACGAG CTCACGGTGC TGGTGTTCCT CACCTTGTCA GTGGTGGTGA TGAAGTTTCT 1380

CCTGGCGGCG CTGGTCCTGT CTCTCATTCT GCCGAGGAGC AGCCAGTACA TCAAGTGGAT 1440

CGTCTCTGCG GGGCTTGCCC AGGTCAGCGA GTTTTCCTTT GTCCTGGGGA GCCGGGCGCG 1500

AAGAGCGGGC GTCATCTCTC GGGAGGTGTA CCTCCTTATA CTGAGTGTGA CCACGCTCAG 1560

CCTCTTGCTC GCCCCGGTGC TGTGGAGAGC TGCAATCACG AGGTGTGTGC CCAGACCGGA 1620

GAGACGGTCC AGCCTCTGAT GGCTCGGAGA TGATGGACCG TGGAAGGGAA GCGTCTGTGG 1680

GGAGTGAGCG CTTAGATGGC CAGCAGCTGC TCCTTCTGGG AAGCTCGCAC CTTGGCAACA 1740

GAACAGCCCT CTAGCAGAGC GTCAGTGCAG TCGTGTTATC CCGGCTTTTA CAGAATATTC 1800

TTGTCCTATT TTAGAATTTT CCGGAGTAGT TTATTTGCAG TCTGTTGATT ATGTGCAGTA 1860

GACCCGGGAC ACTGCGTTTT ACCGATCACC TTGAATGTGG TGCCTGGATG TGCCTTTTTT 1920

TTTTTTCCCT GAAATTATTA TTAATTTTCT ATKGKGAGTT CATCAGTTCA TAGTTTTTTT 1980

AGTAAAGAAG CAAAATTAAA AGGCTTTTAA AAATGTACAA CTTCAGAATT ATAATCTGTT 2040

AGTCAAATAT TTGTTATTAA ACATTTCTGT AATATGAAGT TGTAATCCTG GCCGTGAGCT 2100

TGGAAGCTTA CTTTTGATTC TTAAAGCCTA TGTTTTCTAA AATGAGACAA ATACGGATGT 2160

CTATTTGCCT TTTATTGTAA CTTTTAAATG AAATAATTTC ATGTCAATTT CTATTAGATA 2220

TATCACTTAA AATATTTGGT TTTAAATCAC AAGAATATGT ATTCTTTAAT AAAGATAATT 2280

TATGATCATG GTATAATTAA TTGAAATTTA TTAAAATCTG TTTTTATTAA AAAAAAAAAA 2340

AAA 2343

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 524 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Glu Leu Leu Ala Ala Glu Lys His Gln Val Glu Ala Leu Lys Asn
1 5 10 15

Met Gln His Gln Asn Gln Ser Leu Ser Met Leu Asp Glu Ile Leu Glu

```
                20                  25                  30

Asp Val Arg Lys Ala Ala Asp Arg Leu Glu Glu Glu Ile Glu Glu His
        35                  40                  45

Ala Phe Asp Asp Asn Lys Ser Val Lys Gly Val Asn Phe Glu Ala Val
    50                  55                  60

Leu Arg Val Glu Glu Glu Glu Ala Asn Ser Lys Gln Asn Ile Thr Lys
65                  70                  75                  80

Arg Glu Val Glu Asp Asp Leu Gly Leu Ser Met Leu Ile Asp Ser Gln
                85                  90                  95

Asn Asn Gln Tyr Ile Leu Thr Lys Pro Arg Asp Ser Thr Ile Pro Arg
            100                 105                 110

Ala Asp His His Phe Ile Lys Asp Ile Val Thr Ile Gly Met Leu Ser
        115                 120                 125

Leu Pro Cys Gly Trp Leu Cys Thr Ala Ile Gly Leu Pro Thr Met Phe
    130                 135                 140

Gly Tyr Ile Ile Cys Gly Val Leu Leu Gly Pro Ser Gly Leu Asn Ser
145                 150                 155                 160

Ile Lys Ser Ile Val Gln Val Glu Thr Leu Gly Glu Phe Gly Val Phe
                165                 170                 175

Phe Thr Leu Phe Leu Val Gly Leu Glu Phe Ser Pro Glu Lys Leu Arg
            180                 185                 190

Lys Val Trp Lys Ile Ser Leu Gln Gly Pro Cys Tyr Met Thr Leu Leu
        195                 200                 205

Met Ile Ala Phe Gly Leu Leu Trp Gly His Leu Leu Arg Ile Lys Pro
    210                 215                 220

Thr Gln Ser Val Phe Ile Ser Thr Cys Leu Ser Leu Ser Ser Thr Pro
225                 230                 235                 240

Leu Val Ser Arg Phe Leu Met Gly Ser Ala Arg Gly Asp Lys Glu Gly
                245                 250                 255

Asp Ile Asp Tyr Ser Thr Val Leu Leu Gly Met Leu Val Thr Gln Asp
            260                 265                 270

Val Gln Leu Gly Leu Phe Met Ala Val Met Pro Thr Leu Ile Gln Ala
        275                 280                 285

Gly Ala Ser Ala Ser Ser Ser Ile Val Val Glu Val Leu Arg Ile Leu
    290                 295                 300

Val Leu Ile Gly Gln Ile Leu Phe Ser Leu Ala Ala Val Phe Leu Leu
305                 310                 315                 320
```

Cys Leu Val Ile Lys Lys Tyr Leu Ile Gly Pro Tyr Tyr Arg Lys Leu
325 330 335

His Met Glu Ser Lys Gly Asn Lys Glu Ile Leu Ile Leu Gly Ile Ser
340 345 350

Ala Phe Ile Phe Leu Met Leu Thr Val Thr Glu Leu Leu Asp Val Ser
355 360 365

Met Glu Leu Gly Cys Phe Leu Ala Gly Ala Leu Val Ser Ser Gln Gly
370 375 380

Pro Val Val Thr Glu Glu Ile Ala Thr Ser Ile Glu Pro Ile Arg Asp
385 390 395 400

Phe Leu Ala Ile Val Phe Phe Ala Ser Ile Gly Leu His Val Phe Pro
405 410 415

Thr Phe Val Ala Tyr Glu Leu Thr Val Leu Val Phe Leu Thr Leu Ser
420 425 430

Val Val Val Met Lys Phe Leu Leu Ala Ala Leu Val Leu Ser Leu Ile
435 440 445

Leu Pro Arg Ser Ser Gln Tyr Ile Lys Trp Ile Val Ser Ala Gly Leu
450 455 460

Ala Gln Val Ser Glu Phe Ser Phe Val Leu Gly Ser Arg Ala Arg Arg
465 470 475 480

Ala Gly Val Ile Ser Arg Glu Val Tyr Leu Leu Ile Leu Ser Val Thr
485 490 495

Thr Leu Ser Leu Leu Leu Ala Pro Val Leu Trp Arg Ala Ala Ile Thr
500 505 510

Arg Cys Val Pro Arg Pro Glu Arg Arg Ser Ser Leu
515 520

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 324 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CGCAGCCCGG GCCATGCCGC ACGGCTGCTG ACCGCACGCA GGGGCCGGCC CCGAGGACAC 60

ATGCGGCGGC CTTTGCCGCC TCGCCCCTGA CCCTCTGCCC TGTTCTCCAT GTTGCATTTC 120

TCGTCAGTTT CTCGGGCGGT GTAGCTGCCG CTGCCACCAG AGCCGGCGGG GCATCGCGCT 180

GCTCATTCAT CCGGCCGCAC TTTCTTTTCC GTTTCCACCC ATCCCTTCCC ATTTCCTTCT 240

CCCTTTCCCC GCCAGCTTCG CATCCATCTC CCCCACCCCG TAACCCCTCC TGCCTCCATC 300

CACCGGGGCT ATTGCCGCAA AAGA 324

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 550 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCTGAACATT TCAGAAATAC AGAAGTTGAA GCAGCAGCTT ATGCAGGTAG AGCGGGAAAA 60

GGCCATTCTT TTGGCCAACC TACAGGAGTC ACAGACACAG CTGGAACACA CCAAGGGGGC 120

ACTGACGGAG CAGCATGAGC GGGTGCACCG GCTCACAGAG CACGTCAATG CCATGAGGGG 180

CCTGCAAAGC AGCAAGGAGC TCAAGGCTGA GCTGGACGGG GAGAAGGGCC GGGACTCAGG 240

GGAGGAGGCC CATGACTATG AGGTGGACAT CAATGGTTTA GAGATCCTTG AATGCAAATA 300

CAGGGTGGCA GTAACTGAGG TGATTGATCT GAAAGCTGAA ATTAAGGCCT TAAAGGAGAA 360

ATATAATAAA TCTGTAGAAA ACTACACTGA TGAGAAGGCC AAGTATGAGA GTAAAATCCA 420

GATGTATGAT GAGCAGGTGA CAAGCCTTGA GAAGACCACC AAGGAGAGTG GTGAGAAGAT 480

GGCCCACATG GAGAAGGAGT TGCAAAAGAT GACCAGCATA GCCAACGAAA ATCACAGTAC 540

CCTTAATACG 550

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 170 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Gln Val Glu Arg Glu Lys Ala Ile Leu Leu Ala Asn Leu Gln Glu
1 5 10 15

Ser Gln Thr Gln Leu Glu His Thr Lys Gly Ala Leu Thr Glu Gln His
20 25 30

Glu Arg Val His Arg Leu Thr Glu His Val Asn Ala Met Arg Gly Leu
35 40 45

Gln Ser Ser Lys Glu Leu Lys Ala Glu Leu Asp Gly Glu Lys Gly Arg
50 55 60

Asp Ser Gly Glu Glu Ala His Asp Tyr Glu Val Asp Ile Asn Gly Leu
65 70 75 80

Glu Ile Leu Glu Cys Lys Tyr Arg Val Ala Val Thr Glu Val Ile Asp
85 90 95

Leu Lys Ala Glu Ile Lys Ala Leu Lys Glu Lys Tyr Asn Lys Ser Val
100 105 110

Glu Asn Tyr Thr Asp Glu Lys Ala Lys Tyr Glu Ser Lys Ile Gln Met
115 120 125

Tyr Asp Glu Gln Val Thr Ser Leu Glu Lys Thr Thr Lys Glu Ser Gly
130 135 140

Glu Lys Met Ala His Met Glu Lys Glu Leu Gln Lys Met Thr Ser Ile
145 150 155 160

Ala Asn Glu Asn His Ser Thr Leu Asn Thr
165 170

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 505 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTCCATGAGT GAATTCATCC AAGGGCACGG GTTCAGCAAG GAAAAAAGGT TAACCGTGGT 60

TCCACCAGCA AAAAGAGATT GTCAGCAGCC TGTGCTTCCG TACCGCCACA GTGTTCACAA 120

CTAGCCGGGA GGCAAGACTG CCCAACTGTC AGTCCTGACA CAGCTCTCCC TGAGGAGCAG 180

CCACATTCCA GCTCCCAGTG CGCCCCTCTC CACTGTCTCT CCAAGCCTCC TCACCCCTAG 240

TCTTCATCTC CTGTGGACAA ACATCTGGGG TGGAAGTTTT GTAGCCACAC ACAGGATACT 300

GCCCAAGATC CAGCGGGTGT TTTCTTCTCG GTTGTTAGAT GTACAATTGG ATTAATGTCC 360

ATCGTTTTGG AAGACGAGAG AAAGTTGAGA AGAACACGAA GCACAGACCC TGATGTGATA 420

AAACATTTTG TGGTTTCTCT GAGTCACAGA TAAACTTCTG CCATCAAATG GCTACAGTTC 480

ATTTAAATTT AAAAAAAAAA AAAAA 505

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 481 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGATACTGTA ATAAATAGGA GACAGCTACA GTGATCCAAC TAAACCAACA GGGGATTTTC 60

ATCAGCACTT CCCTGGTGTA ATCATGGTAC AGATTATTAA AGACACGAAT GAATTTAAAA 120

CATTTTTGAC AGCTGCCGGA CACAAACTCG CAGTGGTTCA ATTTTCTTCG AAACGGTGTG 180

GTCCCTGCAA AAGGATGTTT CCTGTTTTCC ATGAGCTGGC TGAAACTTGT CACATCAAAA 240

CAATACCCAC ATTTCAGATG TTCAAGAAAA GCCAGAAGGT AACCCTATTC TCAAGAATCA 300

AAAGAATAAT TTGCTGTTAT AGAAGTGGAT TCATGAGCAA CCTGATTTTT GAGTTTTGTG 360

GAGCCGATGC TAAAAAATTG GAAGCCAAGA CTCAAGAATT AATGTAAGCT GATCTCCAAG 420

GCAAAATACA CTTGTGACAT TTGAAAAGGC AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 480

A 481

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Val Gln Ile Ile Lys Asp Thr Asn Glu Phe Lys Thr Phe Leu Thr
1 5 10 15

Ala Ala Gly His Lys Leu Ala Val Val Gln Phe Ser Ser Lys Arg Cys
20 25 30

Gly Pro Cys Lys Arg Met Phe Pro Val Phe His Glu Leu Ala Glu Thr
35 40 45

Cys His Ile Lys Thr Ile Pro Thr Phe Gln Met Phe Lys Lys Ser Gln
50 55 60

Lys Val Thr Leu Phe Ser Arg Ile Lys Arg Ile Ile Cys Cys Tyr Arg
65 70 75 80

Ser Gly Phe Met Ser Asn Leu Ile Phe Glu Phe Cys Gly Ala Asp Ala
85 90 95

Lys Lys Leu Glu Ala Lys Thr Gln Glu Leu Met
100 105

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1864 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGCCAAAGAG GCCTATTCCT GTGTGCAATC AGTACCTTGA AGGCAGAACA TTCTGAATAA 60

AGTTGGAAAA AGAACAGCTT TGCTTTGCAA AGATTGATGA CAGACTGGTT CCTCAGAGGC 120

CTAGGCTACC CGTCACCCCT TTTTCCAGAG CGAGGGCCTG GAATGAAGGC AGTTTATCCT 180

CTGTCCCTGG AGCCTGGGGT TTGCTTTGGC TCCTTGAGGT GGAAGAGACT AAGAGGGCAG 240

CTGCCCAGAG CAGCTGTGTG TACCTGGCTC CTCTCAGGCT TCCTGATCCC TTCCATTGCA 300

CTGCGCCTTA TCCCTCAGCC AGCCAGACAG CCTCCCTGCT CCTGACCAGC AGATACGTTT 360

CGGAGTGGTT GGTGTGGTTT TTGTGATGAG GGCAGCACGT GGTGGCCAAG GTGGCAAGCT 420

GAGTCTCACA GGCTCACTCC CTCGTTGGTT CCCTGTGGGA ATGGTAGGCC AGGCCCARTA 480

AGCCATGCCC CAACACGTCC TCTCCTCCGG AGGAAGGGCC AGCTGCCARC TGARTCAGCA 540

GCTAGTCCAT AGCACAGCCT TATAACTGTA AAGCCAGGCA TTGCCCATGA GCAGAGCTGG 600

AACCAGAGCT TCAGTCAGTA AGAGGGAGGA TTACCTTCAG GAGAAGGCAA GGAAGAAAAC 660

TGGCTGCTAT CTTTATAGTT CCACTGCCCT AACCAAGTGT CCACATTCTA AATGTGTAGT 720

GTCCATCCCT TATGTAATAG TGGTTTCCCG CCCAAAGTGA GACTTTCCTT TTAATTGGAG 780

AAGGGTATAG AGGTAGTCCA GGTGGGAACG CCAGAAGTGC TGATTGCCCA GCCATTGGGA 840

CCACCTGTTC TTGCCCCACT ACCCTCTAGT GGGAGGCCAA AGTAAAGGCT GGCTGGTGGG 900

TGTCTGTGGA TTGAGGATGT GGCAGGGACT GGTCCTCCCA CCTCCCTCTG GCCAAAGATG 960

GGCTTTGCCC GCTGTGTGCC TGTCACCACC CACCAGCAGT CATGCCCTGG GCTTCCCAAA 1020

TGGAGAGGTA GCAGGCAACG TTTTTAAAAA GAAAGAAAAC AGGAAACTGT ATTGTGTCGG 1080

GGGAGGCGGG AGGGAGATGA GGAAACGGTT TGGATTTTGT GTGTGGGAGG GTATTTTTTG 1140

GGGGTAGTTG TCTGTAACTT TCCTAAGTGC TTTTTTTCCT TTTCTTTTTT AAAGTAAGTT 1200

GCAGGCTTTG GCTTGGAAAA CCCCAGGGGG ATGGGGGGCA GAAACCTGAG GCTGCTGCCC 1260

TTTATCTGCC TTCACGGTAC TGTCCCCTTC CCCCAGCTCC TCCCTGACCC CATGGGCCAG 1320

GCCTCAGACC TTCCAGCTAA CCGCTTCCCA TGAGCCACTA CTCTGATGTC AGCCTATAAC 1380

CAAAGGAGCT GGGGGGTCCA GGCCTGGTGA CCAACCTTTC TCAGCCCACT CAATCAGGGT 1440

GCTCCCCACC TGCAGGCAGG AGGCAACACC CTATCTGCTA CCATCAGCCC CTTCCAGAGC 1500

CCATCTGCCC CGCCCAGCCC TGCCCTGCCC AGCCATACCC TGCTCTGCCC CATCTGGGGG 1560

TGCCCTGCTC AGGGATGGGC TGGCAGGGCT GTACCCAGCC TCCCTGGTAA GCAGAGACTC 1620

AAGAAACCTC TGGGGTCCTG TTTTCTGGTC GTGTGATCCC AGGGGTGCAC ATGGGCCCCT 1680

TGGGTGTCTG AACAGAAGGG CATGGGAGGG AGGGCTGCAC CCCTGCAGTC TTACTCTGCT 1740

GGTGTAGCGG GCAGMTGCCC ACTCCCACCC CACCCTGCAC CGCGGGCTCC TGAGTCGGCA 1800

GATTAAGCAT TTTATAAATT GTATTTTAAA TACATGTTTT AAACTTGTCA AAAAAAAAAA 1860

AAAA 1864

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 102 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Val Leu Pro Thr Cys Arg Gln Glu Ala Thr Pro Tyr Leu Leu Pro Ser
1 5 10 15

Ala Pro Ser Arg Ala His Leu Pro Arg Pro Ala Leu Pro Cys Pro Ala
20 25 30

Ile Pro Cys Ser Ala Pro Ser Gly Gly Ala Leu Leu Arg Asp Gly Leu
35 40 45

Ala Gly Leu Tyr Pro Ala Ser Leu Val Ser Arg Asp Ser Arg Asn Leu
50 55 60

Trp Gly Pro Val Phe Trp Ser Cys Asp Pro Arg Gly Ala His Gly Pro
65 70 75 80

Leu Gly Cys Leu Asn Arg Arg Ala Trp Glu Gly Gly Leu His Pro Cys
85 90 95

Ser Leu Thr Leu Leu Val
100

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1041 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AGCCCTCGTA CTGATTTCCA TCGTTGCATT TACAACTGCT ACAAAAATGC CAGCACTCCA 60

TCGACATGAA GAAGAGAAAT TCTTCTTAAA TGCCAAAGGC CAGAAAGAAA CTTTACCCAG 120

CATATGGGAC TCACCTACCA AACAACTTTC TGTCGTTGTG CCTTCAAACA ATGAAGAAAA 180

ACGGTTGCCT GTGATGATGG ATGAAGCTCT GAGCTATGTA GAGAAGAGAC AGAAACGAGA 240

TCCTGCGTTC ACTTATGAAG TGATAGTAGT TGATGATGGC AGTAAAGATC AGACCTCAAA 300

GGTAGCTTTT AAATATTGCC AGAAATATGG AAGTGACAAA GTACGTGTGA TAACCCTGGT 360

GAAGAATCGT GGAAAAGGTG GAGCGATTAG AATGGGTATA TTCAGTTCTC GAGGAGAAAA 420

GATCCTTATG GCAGATGCTG ATGGAGCCAC AAAGTTTCCA GATGTTGAGA AATTAGAAAA 480

GGGGCTAAAT GATCTACAGC CTTGGCCTAA TCAAATGGCT ATAGCATGTG GATCTCGAGC 540

TCATTTAGAA AAAGAATCAA TTGCTCAGCG TTCTTACTTC CGTACTCTTC TCATGTATGG 600

GTTCCACTTT CTGGTGTGGT TCCTTTGTGT CAAAGGAATC AGGGACACAC AGTGTGGGTT 660

CAAATTATTT ACTCGAGAAG CAGCTTCACG GACGTTTTCA TCTCTACACG TTGAACGATG 720

GGCATTTGAT GTAGAACTAC TGTACATAGC ACAGTTCTTT AAAATTCCAA TAGCAGAAAT 780

TGCTGTCAAC TGGACAGAAA TTGAAGGTTC TAAATTAGTT CCATTCTGGA GCTGGCTACA 840

AATGGGTAAA GACCTACTTT TTATACGACT TCGATATTTG ACTGGTGCCT GGAGGCTTGA 900

GCAAACTCGG AAAATGAATT AGGTTGTTTG CAGTCTTCAG TTGTGTTCTT ATGCTTCAGT 960

GTCACATTTC ATTTCATTTG AAACTAAAAT TTTAAGTAAA GCTGAAATAA ACTTCTTGTC 1020

ATTGTCAAAA AAAAAAAAAA A 1041

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 291 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met Pro Ala Leu His Arg His Glu Glu Glu Lys Phe Phe Leu Asn Ala
1 5 10 15

Lys Gly Gln Lys Glu Thr Leu Pro Ser Ile Trp Asp Ser Pro Thr Lys
20 25 30

Gln Leu Ser Val Val Val Pro Ser Asn Asn Glu Glu Lys Arg Leu Pro
35 40 45

Val Met Met Asp Glu Ala Leu Ser Tyr Val Glu Lys Arg Gln Lys Arg
50 55 60

```
Asp Pro Ala Phe Thr Tyr Glu Val Ile Val Val Asp Asp Gly Ser Lys
65                  70                  75                  80

Asp Gln Thr Ser Lys Val Ala Phe Lys Tyr Cys Gln Lys Tyr Gly Ser
                85                  90                  95

Asp Lys Val Arg Val Ile Thr Leu Val Lys Asn Arg Gly Lys Gly Gly
            100                 105                 110

Ala Ile Arg Met Gly Ile Phe Ser Ser Arg Gly Glu Lys Ile Leu Met
        115                 120                 125

Ala Asp Ala Asp Gly Ala Thr Lys Phe Pro Asp Val Glu Lys Leu Glu
    130                 135                 140

Lys Gly Leu Asn Asp Leu Gln Pro Trp Pro Asn Gln Met Ala Ile Ala
145                 150                 155                 160

Cys Gly Ser Arg Ala His Leu Glu Lys Glu Ser Ile Ala Gln Arg Ser
                165                 170                 175

Tyr Phe Arg Thr Leu Leu Met Tyr Gly Phe His Phe Leu Val Trp Phe
            180                 185                 190

Leu Cys Val Lys Gly Ile Arg Asp Thr Gln Cys Gly Phe Lys Leu Phe
        195                 200                 205

Thr Arg Glu Ala Ala Ser Arg Thr Phe Ser Ser Leu His Val Glu Arg
    210                 215                 220

Trp Ala Phe Asp Val Glu Leu Leu Tyr Ile Ala Gln Phe Phe Lys Ile
225                 230                 235                 240

Pro Ile Ala Glu Ile Ala Val Asn Trp Thr Glu Ile Glu Gly Ser Lys
                245                 250                 255

Leu Val Pro Phe Trp Ser Trp Leu Gln Met Gly Lys Asp Leu Leu Phe
            260                 265                 270

Ile Arg Leu Arg Tyr Leu Thr Gly Ala Trp Arg Leu Glu Gln Thr Arg
        275                 280                 285

Lys Met Asn
    290
```

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CNCCATCGGG GAACACCAGA AAGAACACT 29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TNTCTGGCAT ATCCGTCAGG TTAAACTCC 29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CNCTGGTTCT ACATCAATAC CAGCTTTCC 29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TNACAACAGT GATATTTGAG AGCTTCAAG 29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CNGTAACACC TCTCCAACGC TTTCGATGC 29

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GNCAAGGACA GACACGTGGA AATGAAGAC 29

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ANGTCCACCT CATAGTCATG GGCCTCCTC 29

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TNTCAGCCAG CTCATGGAAA ACAGGAAAC 29

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CNTGGGAAGC GGTTAGCTGG AAGGTCTGA 29

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

TNTCTTCTTC ATGTCGATGG AGTGCTGGC 29

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 359 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Arg Val Lys Val Gln Leu Ala Leu Val Phe Phe Lys Asn Leu Ala Ser
1 5 10 15

Ser Cys Thr Leu Asp Ser Ala Leu Asn Ala Ala Ser Tyr Tyr Asn Phe
20 25 30

Thr Val Leu Lys Val Pro Arg Ser Met Thr Asp Pro Gln Asn Met Glu
35 40 45

Phe Gln Val Pro Val Ile Leu Thr Ser Gln Ala Asn Ala Pro Leu Leu
50 55 60

Ala Gly Asn Thr Cys Gln Asn Val Val Ser Gln Val Thr Tyr Glu Ile
65 70 75 80

Glu Thr Asn Gly Thr Phe Gly Ile Gln Lys Val Ser Val Ser Leu Gly
85 90 95

Gln Thr Asn Leu Thr Val Glu Pro Gly Ala Ser Leu Gln Gln His Phe
100 105 110

Ile Leu Arg Phe Arg Ala Phe Gln Gln Ser Thr Ala Ala Ser Leu Thr
115 120 125

Ser Pro Arg Ser Gly Asn Pro Gly Tyr Ile Val Gly Lys Pro Leu Leu
130 135 140

Ala Leu Thr Asp Asp Ile Ser Tyr Ser Met Thr Leu Leu Gln Ser Gln
145 150 155 160

Gly Asn Gly Ser Cys Ser Val Lys Arg His Glu Val Gln Phe Gly Val
165 170 175

Asn Ala Ile Ser Gly Cys Lys Leu Arg Leu Lys Lys Ala Asp Cys Ser
180 185 190

His Leu Gln Gln Glu Ile Tyr Gln Thr Leu His Gly Arg Pro Arg Pro
195 200 205

```
Glu Tyr Val Ala Ile Phe Gly Asn Ala Asp Pro Ala Gln Lys Gly Gly
    210                 215                 220

Trp Thr Arg Ile Leu Asn Arg His Cys Ser Ile Ser Ala Ile Asn Cys
225                 230                 235                 240

Thr Ser Cys Cys Leu Ile Pro Val Ser Leu Glu Ile Gln Val Leu Trp
                245                 250                 255

Ala Tyr Val Gly Leu Leu Ser Asn Pro Gln Ala His Val Ser Gly Val
            260                 265                 270

Arg Phe Leu Tyr Gln Cys Gln Ser Ile Gln Asp Ser Gln Gln Val Thr
        275                 280                 285

Glu Val Ser Leu Thr Thr Leu Val Asn Phe Val Asp Ile Thr Gln Lys
    290                 295                 300

Pro Gln Pro Pro Arg Gly Gln Pro Lys Met Asp Trp Lys Trp Pro Phe
305                 310                 315                 320

Asp Phe Phe Pro Phe Lys Val Ala Phe Ser Arg Gly Val Phe Ser Gln
                325                 330                 335

Lys Cys Ser Val Ser Pro Ile Leu Ile Leu Cys Leu Leu Glu Leu Gly
            340                 345                 350

Val Leu Asn Leu Glu Thr Met
        355
```

What is claimed is:

1. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 12 to nucleotide 800;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 78 to nucleotide 800;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1 to nucleotide 547;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bh389_11 deposited under accession number ATCC 98451;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bh389_11 deposited under accession number ATCC 98451;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bh389_11 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bh389_11 deposited under accession number ATCC 98451;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:2;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

2. The polynucleotide of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with the polynucleotide of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.

5. A process for producing a protein encoded by the polynucleotide of claim 2, which process comprises:

(a) growing a culture of the host cell of claim 3 in a suitable culture medium; and

(b) purifying said protein from the culture.

6. A protein produced according to the process of claim 5.

7. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 178;

(c) fragments of the amino acid sequence of SEQ ID NO:2 comprising eight consecutive amino acids of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone bh389_11 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins.

8. The protein of claim 7, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.

9. The protein of claim 7, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 178.

10. A composition comprising the protein of claim 7 and a pharmaceutically acceptable carrier.

11. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.

12. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 100 to nucleotide 882;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 635 to nucleotide 867;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bk112_15 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bk112_15 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bk112_15 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bk112_15 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

13. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 200 to amino acid 256;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising eight consecutive amino acids of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone bk112_15 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins.

14. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3.

15. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 245 to nucleotide 520;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 181 to nucleotide 527;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bk200_13 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bk200_13 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bk200_13 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bk200_13 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:6;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

16. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) fragments of the amino acid sequence of SEQ ID NO:6 comprising eight consecutive amino acids of SEQ ID NO:6; and

(c) the amino acid sequence encoded by the cDNA insert of clone bk200_13 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins.

17. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5.

18. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 365 to nucleotide 784;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 518 to nucleotide 784;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone di386_3 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone di386_3 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone di386_3 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone di386_3 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:8;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

19. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 140;

(c) fragments of the amino acid sequence of SEQ ID NO:8 comprising eight consecutive amino acids of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone di386_3 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins.

20. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7 and SEQ ID NO:9.

21. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 191 to nucleotide 781;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 56 to nucleotide 492;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone em397_2 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone em397_2 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone em397_2 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone em397_2 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:11;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

22. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:11;

(b) the amino acid sequence of SEQ ID NO:11 from amino acid 1 to amino acid 101;

(c) fragments of the amino acid sequence of SEQ ID NO:11 comprising eight consecutive amino acids of SEQ ID NO:11; and

(d) the amino acid sequence encoded by the cDNA insert of clone em397_2 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins.

23. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:10.

24. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 65 to nucleotide 1636;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 482 to nucleotide 1636;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 487 to nucleotide 1006;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fh170_7 deposited under accession number ATCC 98451;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fh170_7 deposited under accession number ATCC 98451;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fh170_7 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fh170_7 deposited under accession number ATCC 98451;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:13;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:13 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:13;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

25. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:13;

(b) the amino acid sequence of SEQ ID NO:13 from amino acid 142 to amino acid 314;

(c) fragments of the amino acid sequence of SEQ ID NO:13 comprising eight consecutive amino acids of SEQ ID NO:13; and

(d) the amino acid sequence encoded by the cDNA insert of clone fh170_7 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins.

26. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:12.

27. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 41 to nucleotide 550;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fn53_4 deposited under accession number ATCC 98451;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fn53_4 deposited under accession number ATCC 98451;

(e) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fn53_4 deposited under accession number ATCC 98451;

(f) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fn53_4 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:16;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(h).

28. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 40 to amino acid 170;

(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising eight consecutive amino acids of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone fn53_4 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins.

29. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:15, SEQ ID NO:14, and SEQ ID NO:17 .

30. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 84 to nucleotide 404;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 78 to nucleotide 493;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fq505_4 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fq505_4 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fq505_4 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fq505_4 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:19;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:19;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

31. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:19;

(b) the amino acid sequence of SEQ ID NO:19 from amino acid 23 to amino acid 107;

(c) fragments of the amino acid sequence of SEQ ID NO:19 comprising eight consecutive amino acids of SEQ ID NO:19; and

(d) the amino acid sequence encoded by the cDNA insert of clone fq505_4 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins.

32. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:18.

33. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 1439 to nucleotide 1744;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 1241 to nucleotide 1754;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fw13_9 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fw13_9 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fw13_9 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fw13_9 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:21;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

34. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:21;

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 57;

(c) fragments of the amino acid sequence of SEQ ID NO:21 comprising eight consecutive amino acids of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone fw13_9 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins.

35. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:20.

36. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 47 to nucleotide 919;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 124 to nucleotide 452;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone gg619_2 deposited under accession number ATCC 98451;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone gg619_2 deposited under accession number ATCC 98451;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone gg619_2 deposited under accession number ATCC 98451;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone gg619_2 deposited under accession number ATCC 98451;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:23;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

37. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:23;

(b) the amino acid sequence of SEQ ID NO:23 from amino acid 27 to amino acid 135;

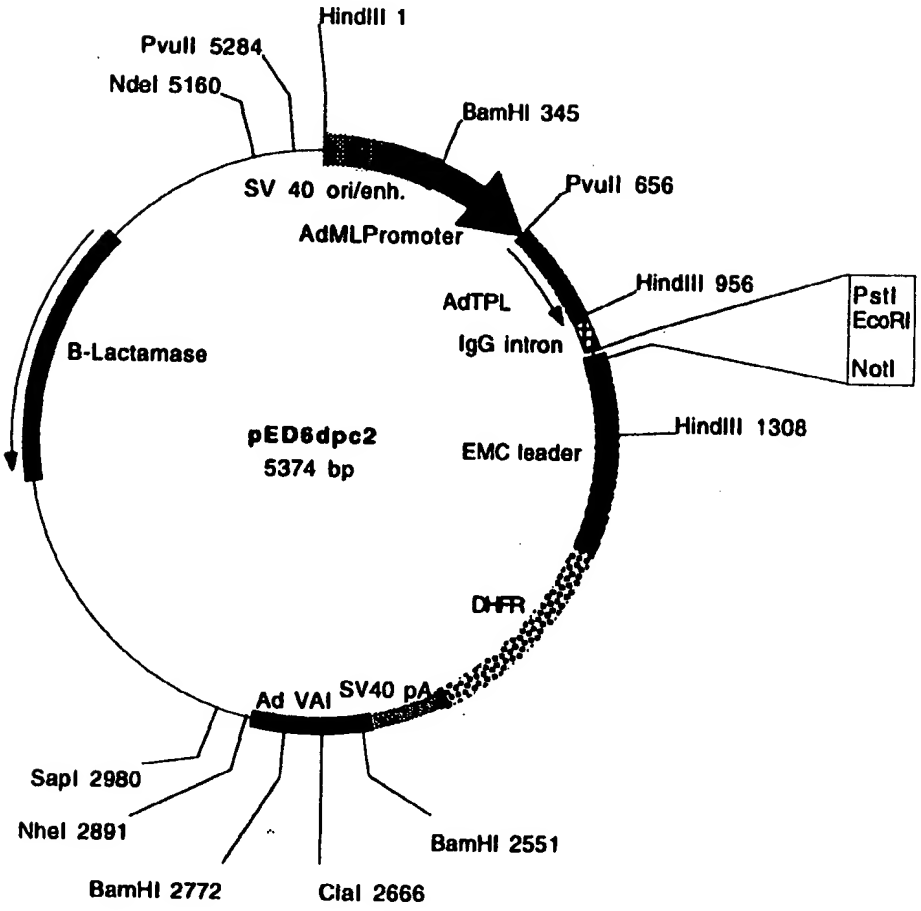
(c) fragments of the amino acid sequence of SEQ ID NO:23 comprising eight consecutive amino acids of SEQ ID NO:23; and

(d) the amino acid sequence encoded by the cDNA insert of clone gg619_2 deposited under accession number ATCC 98451;

the protein being substantially free from other mammalian proteins.

38. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:22.

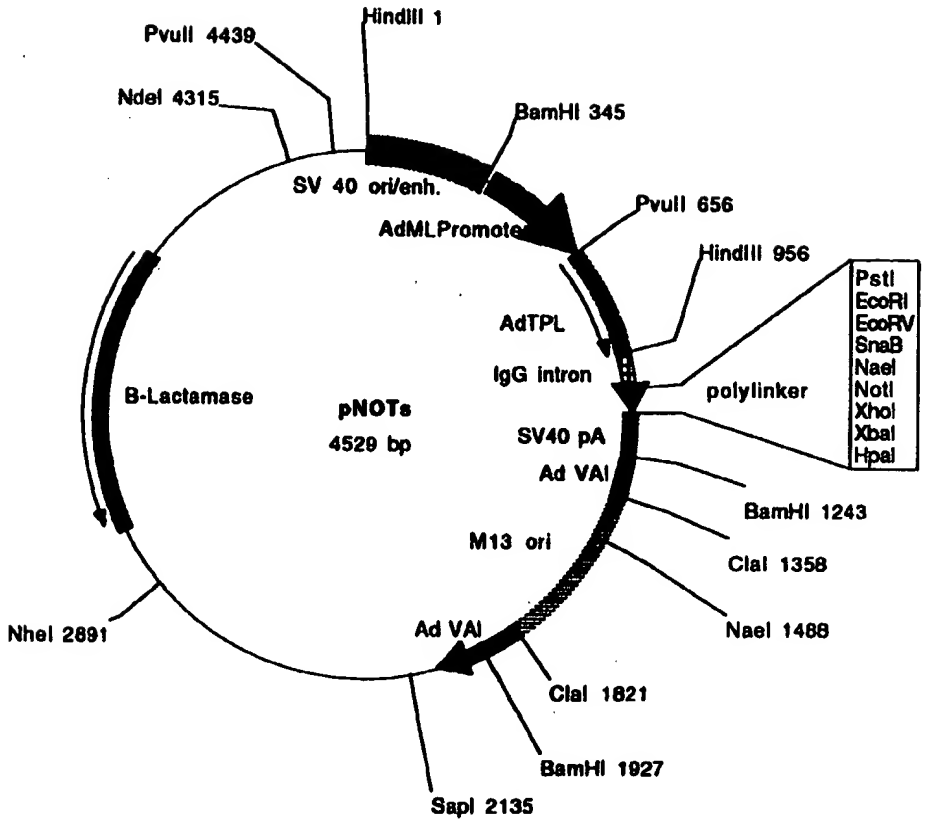
FIGURE 1A



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FIGURE 1B



Plasmid name: pNOTs
Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al,1989. Mol.Cell.Biol.9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of r plication was inserted in th ClaI site. SST cDNAs are cl ned between EcoRI and NotI